# INTERNATIONAL ORGANIZATION FOR STANDARDIZATION

# ORGANISATION INTERNATIONALE DE NORMALISATION

# ISO/IEC/JTC 1/SC 29/WG 11

# CODING OF MOVING PICTURES AND AUDIO

**ISO/IEC JTC 1/SC 29/WG 11 N17698**

**July 2018 – Ljubljana, SI**

|  |  |  |
| --- | --- | --- |
| Source: | Convenor of MPEG  |  |
| Status: | Approved by WG11 |
| Subject: | MPEG Press Release |
| Date: | 20 July 2018 |

**MPEG issues Call for Evidence on Compressed Representation of Neural Networks**

Ljubljana, SI – The 123rd MPEG meeting was held in Ljubljana, SI, from 16 – 20 July 2018

**MPEG issues Call for Evidence on Compressed Representation of Neural Networks**

Artificial neural networks have been adopted for a broad range of tasks in multimedia analysis and processing, media coding, data analytics, translation and many other fields. Their recent success is based on the feasibility of processing much larger and complex neural networks (deep neural networks, DNNs) than in the past, and the availability of large-scale training data sets. As a consequence, trained neural networks contain a large number of parameters (weights), resulting in a quite large size (e.g., several hundred MBs). Many applications require the deployment of a particular trained network instance, potentially to a larger number of devices, which may have limitations in terms of processing power and memory (e.g., mobile devices or smart cameras). Any use case, in which a trained neural network (and its updates) needs to be deployed to a number of devices could thus benefit from a standard for the compressed representation of neural networks.

At its 123rd meeting, MPEG has issued a Call for Evidence (CfE) for compression technology for neural networks. The compression technology will be evaluated in terms of compression efficiency, runtime, and memory consumption and the impact on performance in three use cases: visual object classification, visual feature extraction (as used in MPEG Compact Descriptors for Visual Analysis) and filters for video coding. Responses to the CfE will be analyzed on the weekend prior to and during the 124th MPEG meeting in October 2018 (Macau, CN).

**Network-Based Media Processing – MPEG evaluates responses to call for proposal and kicks off its technical work**

Recent developments in multimedia have brought significant innovation and disruption to the way multimedia content is created and consumed. At its 123rd meeting, MPEG analyzed the technologies submitted by eight industry leaders as responses to the Call for Proposals (CfP) for Network-Based Media Processing (NBMP, MPEG-I Part 8). These technologies address advanced media processing use cases such as network stitching for virtual reality (VR) services, super-resolution for enhanced visual quality, transcoding by a mobile edge cloud, or viewport extraction for 360-degree video within the network environment. NBMP allows service providers and end users to describe media processing operations that are to be performed by the entities in the networks. NBMP will describe the composition of network-based media processing services out of a set of NBMP functions and makes these NBMP services accessible through Application Programming Interfaces (APIs).

NBMP will support the existing delivery methods such as streaming, file delivery, push-based progressive download, hybrid delivery, and multipath delivery within heterogeneous network environments. MPEG issued a Call for Proposal (CfP) seeking technologies that allow end-user devices, which are limited in processing capabilities and power consumption, to offload certain kinds of processing to the network.

After a formal evaluation of submissions, MPEG selected three technologies as starting points for the *(i)* workflow, *(ii)* metadata, and *(iii)* interfaces for static and dynamically acquired NBMP. A key conclusion of the evaluation was that NBMP can significantly improve the performance and efficiency of the cloud infrastructure and media processing services.

**MPEG finalizes 1st edition of Technical Report on Architectures for Immersive Media**

At its 123nd meeting, MPEG finalized the first edition of its Technical Report (TR) on Architectures for Immersive Media. This report constitutes the first part of the MPEG-I standard for the coded representation of immersive media and introduces the eight MPEG-I parts currently under specification in MPEG. In particular, it addresses *three Degrees of Freedom* (3DoF; three rotational and un-limited movements around the X, Y and Z axes (respectively pitch, yaw and roll)), *3DoF+* (3DoF with additional limited translational movements (typically, head movements) along X, Y and Z axes), and *6DoF* (3DoF with full translational movements along X, Y and Z axes) experiences but it mostly focuses on 3DoF. Future versions are expected to cover aspects beyond 3DoF. The report documents use cases and defines architectural views on elements that contribute to an overall immersive experience. Finally, the report also includes quality considerations for immersive services and introduces minimum requirements as well as objectives for a high-quality immersive media experience.

**MPEG releases software for MPEG-I visual activities**

MPEG-I visual is an activity that addresses the specific requirements of immersive visual media for six degrees of freedom virtual walkthroughs with correct motion parallax within a bounded volume. MPEG-I visual covers application scenarios from 3DoF+ with slight body and head movements in a sitting position to 6DoF allowing some walking steps from a central position. At the 123nd MPEG meeting, an important progress has been achieved in software development. A new Reference View Synthesizer (RVS 2.0) has been released for 3DoF+, allowing to synthesize virtual viewpoints from an unlimited number of input views. RVS integrates code bases from Universite Libre de Bruxelles and Philips, who acted as software coordinator. A Weighted-to-Spherically-uniform PSNR (WS-PSNR) software utility, essential to 3DoF+ and 6DoF activities, has been developed by Zhejiang University. WS-PSNR is a full reference objective quality metric for all flavors of omnidirectional video. RVS and WS-PSNR are essential software tools for the upcoming Call for Proposals on 3DoF+ expected to be released at the 124th MPEG meeting in October 2018 (Macau, CN).

**MPEG enhances ISO Base Media File Format (ISOBMFF) with new features**

At the 123rd MPEG meeting, a couple of new amendments related to ISOBMFF has reached the first milestone. Amendment 2 to ISO/IEC 14496-12 6th edition will add the option to have relative addressing as an alternative to offset addressing, which in some environments and workflows can simplify the handling of files and will allow creation of derived visual tracks using items and samples in other tracks with some transformation, for example rotation. Another amendment reached its first milestone is the first amendment to ISO/IEC 23001-7 3rd edition. It will allow use of multiple keys to a single sample and scramble some parts of AVC or HEVC video bitstreams without breaking conformance to the existing decoders. That is, the bitstream will be decodable by existing decoders, but some parts of the video will be scrambled. It is expected that these amendments will reach the final milestone in Q3 2019.

**How to contact MPEG, learn more, and find other MPEG facts**

To learn about [MPEG basics](http://mpeg.chiariglione.org/mpeg-basics), discover [how to participate](http://mpeg.chiariglione.org/who-we-are) in the committee, or find out more about the array of technologies developed or currently under development by MPEG, visit MPEG’s home page at <https://mpeg.chiariglione.org/>. There you will find information publicly available from MPEG experts past and present including tutorials, white papers, vision documents, and requirements under consideration for new standards efforts. You can also find useful information in many public documents by using the search window including publicly available output documents of each meeting (note: some may have editing periods and in case of questions please contact Dr. Christian Timmerer).

Examples of tutorials that can be found there include tutorials for: High Efficiency Video Coding, Advanced Audio Coding, Universal Speech and Audio Coding, and DASH to name a few. A rich repository of white papers can also be found and continues to grow. You can find these papers and tutorials for many of [MPEG’s standards](http://mpeg.chiariglione.org/standards) freely available. Press releases from previous MPEG meetings are also available.

Journalists that wish to receive MPEG Press Releases by email should contact Dr. Christian Timmerer at christian.timmerer@itec.uni-klu.ac.at or christian.timmerer@bitmovin.com or subscribe via <https://lists.aau.at/mailman/listinfo/mpeg-pr>. For timely updates follow us on Twitter (<https://twitter.com/mpeggroup>).

**Further Information**

Future MPEG meetings are planned as follows:

No. 124, Macau SAR, CN, 08 – 12 October 2018

No. 125, Marrakech, MA, 14 – 18 January 2019

No. 126, Geneva, CH, 18 – 22 March 2019

No. 127, Gothenburg, SE, 08 – 12 July 2019

For further information about MPEG, please contact:

Dr. Leonardo Chiariglione (Convenor of MPEG, Italy)

Via Borgionera, 103

I-10040 Villar Dora (TO), Italy

Tel: +39 011 935 04 61

leonardo@chiariglione.org

or

Priv.-Doz. Dr. Christian Timmerer

Alpen-Adria-Universität Klagenfurt | Bitmovin Inc.

9020 Klagenfurt am Wörthersee, Austria, Europe

Tel: +43 463 2700 3621

Email: christian.timmerer@itec.aau.at | christian.timmerer@bitmovin.com